Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 4, 2026

ENHANCING SMART GRID RELIABILITY THROUGH MACHINE LEARNING-BASED ANOMALY DETECTION

[1]Junaid Ur Rahman, [2]Syed Muhammad Ubaid Ur Rahman, [3]Mohd Sumer, [4]Ali Taqi

[1]*University of Roehampton, London*
[2]*University of Roehampton, London*
[3]*University of Roehampton, London*
[4]*Lecturer & PhD Scholar, Department of Computer Science & Information Technology, Govt. College Women University, Sialkot*
[1]junaidurrahman362@gmail.com, [2]syedmuhammadubaidurrahman@gmail.com [3]mrmohdsumer@gmail.com
[4]ali.taqi@gcwus.edu.pk/alitaqi882@gmail.com

DOI: https://doi.org/ https://doi.org/10.5281/zenodo.19951125

Article History
Smart Grid, Machine Learning, Anomaly Detection, Reliability, Predictive Maintenance, Cybersecurity, SAIDI, SAIFI, Deep Learning, Explainable AI

Article History
Received on 25 March, 2026
Accepted on 29 April, 2026
Published on 30 April, 2026

Corresponding Author: *

Abstract

Background: Advanced monitoring and control have been added by the development of smart grids, but they have also raised the complexity and susceptibility of systems to anomalies like faults, cyber-attacks, and data inconsistencies. Conventional rule-based anomaly detectors are becoming less and less sufficient to sustain high volume, real-time data, and thus more intelligent and adaptive solutions are required. Objective: This paper will evaluate the performance of machine learning (ML) to enhance the reliability of smart grids, through anomaly detection, and identify the key issues, performance standards, and deploying barriers, from the perspective of industry practitioners. Methodology: A quantitative survey-based approach was adopted, with a sample of 250 professionals including grid operators, data scientists, academicians, and policy makers. The survey aimed to gather information on current practices, required ML approaches, challenges in deployment and anticipated performance levels. Data analysis was performed using descriptive statistics (mean, standard deviation, frequency, percentage). Results: The results indicate that there is a high agreement that the current traditional approaches cannot work, mainly because of the growing complexity of data and the presence of anomalies that cannot be identified. The most preferred methods are supervised, unsupervised, and hybrid ML techniques. The high performance expectations were established, and the majority of respondents are expected to be precise and recall with an accuracy rate of above 90 percent and detection latency of less than 10 seconds. The main difficulties are a shortage of labeled data, class imbalance, cybersecurity issues, and compatibility with the legacy systems. Nonetheless, despite these obstacles, ML is seen to have a strong beneficial influence on predictive maintenance, fault minimization, and cyber-attack control. Moreover, most respondents demonstrated a favorable attitude to using ML in the future.Conclusion: Machine learning is a revolutionary chance to improve the reliability of smart grids with the help of precise and real-time detection of anomalies.

Nevertheless, to achieve a successful large scale implementation, one has to deal with data limitations, enhance model interpretability, enhance infrastructure integration and acquire skilled human resources.

Introduction

The rapid evolution of smart grids in modern power systems has fundamentally transformed the generation, transmission and use of power. Smart grids, unlike conventional power grids, employ smart communication technologies, Internet of Things (IoT) technologies, smart meters and phasor measurement units with real-time monitoring and control [1]. This transformation has led to more efficient operation and increased system transparency, but it has also posed new challenges, especially with system reliability and resilience [2].

Anomaly detection is a critical problem in smart grids. Anomaly can be due to equipment malfunction, cyber-attack, data inconsistencies or outages [3]. Traditional anomaly detection approaches, which are typically based on rule-based or threshold-based features, are not sufficiently meeting demands of today's smart grids in terms of speed, volume and complexity [4]. One of the challenges with the increasing data velocity and volume is that it may exceed the capacity of conventional monitoring systems, which may result in delayed or missed detection of events.

Machine learning is a potential solution to these challenges. Machine learning models can also identify patterns, learn from the system and detect anomalies more effectively than traditional systems due to the data-driven approach used [5]. This is particularly applicable for detection of non-rare, multi-dimensional and dynamic anomalies [6]. Unlike rule-based systems, machine learning models can adapt to changing grid conditions, improve the accuracy of anomaly detection, and reduce false alarms [7].

Use of machine learning for anomaly detection in smart grids is advantageous [8]. It enables processing of noisy, high-dimensional data, real-time decision making and predictive capability [9]. This is highly valuable in areas such as predictive maintenance, cyber-security and fault detection [10]. However, the use of machine learning in smart grids that are currently in operation is in its infancy despite these advantages [11].

There are many companies at pilot or research stages, so widespread roll-out is still to come. There are several challenges for the large-scale deployment of machine learning-based anomaly systems [12]. These include the lack of labeled anomaly data, class imbalance, and difficulty with integration into the existing systems, along with a lack of interpretability and trust in the models [13]. Furthermore, real-time processing has strict latency requirements, making real-time operation even more difficult.

Reliability is a critical performance metric in smart grids, and it is usually measured by such indices as SAIDI and SAIFI [14]. To ensure service quality and customer satisfaction, the short outage and fast response is needed [15]. The methods based on machine learning can improve these indices considerably as they can identify and react to the faults and cyber attacks with a higher speed [16].

Finally, the anomaly detection using machine learning is one of the research and development challenges in smart grid reliability. While it is a technology with considerable benefits, its implementation will require addressing technical, operational and organisational issues. The proposed research will explore these challenges and prospects by investigating the perceptions, practices and expectations of experts in this field, and thereby advancing our understanding of the use of machine learning in modern power systems.

Problem Statement

Smart grid systems generate large volumes of real-time data through a wide range of sources, including smart meters, sensors and monitoring equipment. The existing rule-based anomaly detection methods are increasingly struggling to deal with such complexity, which results in unrecognised faults, delays in response time and reduced reliability of power systems. Cyber attacks, component malfunction, and data inconsistencies also play a role. Despite the great potential of machine learning in anomaly detection, it is not

implemented extensively due to challenges associated with lack of labelled data, class bias, integration, interpretability and real-time performance. Thus, the impact, issues and implications of machine learning-based anomaly detection for enhancing the reliability of smart grids are highly valuable and the perspective of industry experts on the process of implementation and the future potential should be understood.

Literature Review

Smart Grid Evolution and Reliability Challenges

The transition from conventional power systems to smart grids brings new opportunities for monitoring and control [17]. However, this transformation has also resulted in more complex system usage, making reliability more difficult. The previous systems were more likely to be geared towards a fixed and centralised operation, but the smart grids are intended to be dynamic and distributed [18]. This has made anomaly detection and response more difficult and more sophisticated techniques are required.

Limitations of Traditional Anomaly Detection Methods

Current anomaly detection methods are rule-based. These approaches are easy to understand but not flexible and don't identify complex patterns [19]. These techniques are not able to cope with the vast amount of data generated by the growing number of IoT devices and other high-frequency data streams [20]. It may result in either false alarms or missed alarms, causing confusion.

Machine Learning in Anomaly Detection

Machine learning has been significant as it deals with complex data and can find hidden patterns [21]. Supervised learning techniques, such as support vector machines and random forests, have been frequently used for classification [22]. But they require labelled data which is not always available in anomaly detection.

However, unsupervised learning techniques, such as clustering and autoencoders, overcome this limitation by identifying anomaly without using labelled data [23]. These approaches are applicable in smart grids where anomalies are not frequent and vary in nature [24]. The latest deep learning methods with long short-term memory networks and convolutional neural networks go a step further by extracting temporal and spatial features of the data [25].

Ensemble and hybrid techniques are approaches that combine multiple techniques to improve their performance and reliability [26]. This method employs the advantages of multiple algorithms to achieve improved accuracy and generalization. In addition, the recent approaches such as graph neural networks can perform graph-based analysis, which can be very useful in power systems.

Challenges in Machine Learning Deployment

Despite the advantages of machine learning, there are several issues with its implementation. Lack of anomalies data is one of the key issues which limited the supervised learning performance [27]. Another significant problem is the imbalance in classes, because anomalies are a minor subset of the data.

Another problem is in the interpretability of the models, particularly to grid operators who need to have explainable and transparent decision-making processes. Skepticism and refusal to adopt can be caused by black boxes [28]. In addition, the technical difficulty is in integrating machine learning applications and current technologies, such as SCADA and energy management systems.

There are also time processing requirements that add to the complexity. Field operation in smart grids requires real-time responses and detections (typically in seconds). The need is for low latency and high accuracy [29]. Finally, machine learning pipeline security concerns have to be considered to avoid adversarial attacks and spoofing.

Impact on Smart Grid Reliability

Anomaly detection using machine learning can greatly improve the reliability of smart grids. It allows the rapid identification of equipment problems, minimizing downtime and maintenance. It is also capable of enhancing detecting cyber threats and maintaining system security and resilience. Machine learning has made predictive maintenance possible, which can be used to implement proactive interventions and reduce disruptions [30].

In addition, highly developed anomaly detection systems can be used to facilitate self-healing grid processes, in which automated solutions address

the effect of faults [31]. This will result in better reliability measures and overall system performance.

Future Research Directions

The development of research in this field in the future is aimed at eliminating the challenges that exist and enhancing the performance of the models. The production of synthetic labeled datasets is one of the top priorities in order to address the lack of data. There should also be the development of explainable machine learning models to make it more trusted and adopted by the operators.

The other research topics are the real-time validation, light-weight edge device models and robust detection models against adversaries. These advancements will play a significant role in enabling the adoption of machine learning in smart grids as the field progresses.

Research Questions

- To what extent are traditional anomaly detection methods considered insufficient in smart grid systems?
- How significantly do undetected anomalies impact smart grid reliability?
- Which machine learning techniques are most preferred for anomaly detection in smart grids?
- What are the major challenges faced in deploying machine learning-based anomaly detection systems?
- What performance requirements (precision, recall, latency) are expected for effective anomaly detection?
- What are the key barriers to adoption of machine learning in smart grid environments?
- How does machine learning contribute to improving reliability metrics such as fault detection and cyber-attack response?
- What are the future priorities for enhancing anomaly detection using machine learning?

Research Objectives

- To analyze the limitations of existing anomaly detection methods in smart grids
- To evaluate the role of machine learning in improving anomaly detection accuracy and efficiency
- To identify the most suitable machine learning techniques for smart grid applications
- To examine the key technical and organizational challenges in implementation
- To assess expected performance standards for real-time anomaly detection
- To explore barriers affecting adoption and deployment
- To determine the perceived impact of machine learning on smart grid reliability
- To identify future research and development priorities in this domain

Methodology

Research Design

The aim of this paper is to explore quantitative, descriptive type of research in order to understand the potential of machine learning to improve the reliability of smart grid. The approach is based on the collection of structured responses of experts to assess the views, challenges and expectations regarding ML-based anomaly detection. A questionnaire was preferred to achieve systematic data gathering and statistical taxonomy to determine the trends and patterns.

Data Collection Method

Primary data was collected using a questionnaire, which was developed after literature review and in line with the research goals. To get the current practices, preferred ML approaches, performance expectations and barriers in implementation, the questionnaire included a set of Likert scale responses, multiple choice and categorical responses.

The questionnaire was made available online for easy access in terms of inclusion of various professions of the smart grid sector.

Sampling Technique and Respondents

A purposive sampling technique was employed to target individuals with relevant expertise in smart grid operations and machine learning. The respondents included:

- Grid operators and utility engineers
- Data scientists and ML specialists
- Academic researchers
- Solution providers and vendors
- Regulatory and policy advisors

A total of 250 valid responses were collected, ensuring a diverse and balanced representation of technical and operational perspectives.

Data Analysis Techniques

The collected data was analyzed using descriptive statistical methods, including:

- Mean and standard deviation (for Likert-scale responses)
- Frequency and percentage distributions (for categorical variables)

These techniques enabled the identification of trends, consensus levels, and variability among respondents. The analysis was structured around key research dimensions such as current state, ML preferences, challenges, performance expectations, and perceived impact.

Measurement Variables

The study evaluated multiple variables, including:

- Effectiveness of traditional methods
- Preferred ML techniques
- Deployment challenges (e.g., data scarcity, integration, cybersecurity)
- Performance metrics (precision, recall, latency)
- Adoption barriers and R&D priorities
- Perceived impact on reliability indicators (e.g., SAIDI, SAIFI)

These variables were measured using standardized scales to ensure consistency and comparability.

Ethical Considerations

This research study has been carried out voluntarily and respondents were informed of the purpose of the study. This ensured anonymity and confidentiality of the data and the responses was to be used only for educational and research purposes.

Limitations of the Methodology

While the study can provide valuable insights, it is limited by using self-reported data which can be bias. And purposive sampling can limit the capacity to generalise the results to all smart grid stakeholders. However, the availability of skilled personnel offset these prohibitions and provide informed responses.

Results of the Study

The Findings of the Study are the objective statements of what was found in the data collection and analysis and they address the research questions or hypotheses. This section generates reports of facts, such as statistical descriptions, patterns or effects but it does not interpret or discuss their broader significance. It provides the evidence on which conclusions are drawn.

Table 1: *Respondent Profile*

Characteristic	Category	Count	Percentage
Primary Role	Grid operator / utility engineer	85	34.0%
	Data scientist / ML specialist	60	24.0%
	Academic researcher	45	18.0%
	Solution provider / vendor	30	12.0%
	Regulatory / policy advisor	20	8.0%
	Other	10	4.0%
Experience	< 2 years	30	12.0%
	2–5 years	70	28.0%
	6–10 years	90	36.0%
	> 10 years	60	24.0%
Org. ML Involvement	No involvement	40	16.0%
	Pilot / research stage	110	44.0%
	Partial deployment	70	28.0%
	Full production	30	12.0%

As for primary role, grid operators and utility engineers (34%) dominate the dataset, which indicates the operational nature of the data. This is reflected in the large number of data scientists/Machine learning experts (24%) which suggests a high level of data science expertise. Researchers (18%) also contribute research-backed viewpoints and vendors (solution providers) (12%) and regulatory/policy advisors (8%) cover implementation and policy issues. The "other" (4%) category does not influence trends. The mix of roles, therefore, provides a balanced group of

practitioners and experts with increased trust in the information on operational and innovation matters.

The respondents are mostly middle career with 36% having between 6-10 years career and 28% between 2-5 years. This indicates a sample of persons that have enough experience to give reliable answers. The mature respondents (over 10 years) are 24% which would add to the maturity and a strategy-oriented survey and the inexperienced (less than 2 years) 12% which reduces the chance of inexperience in the responses. The mixture of experienced people indicates a robust sample that could consider the current state of the practices and technologies.

Regarding organizational ML involvement, the majority of respondents state that it takes place at the pilot or research level (44%), which means that the utilization of machine learning in most companies remains in an early stage of development. Some 28% report partial deployment, indicating a slow integration toward operations. Nevertheless, 12% of them are offering full production-level deployment which emphasizes a comparatively low degree of maturity in the deployment of ML. In the meantime, 16% state that there is no involvement, indicating that a significant number of organizations have not started to adopt ML.

Overall, the data demonstrates that although the level of expertise and experience is high, organizational adoption of machine learning in the smart grid remains mostly in a transitional phase, with the potential to operate on pilot projects to full deployment being immense.

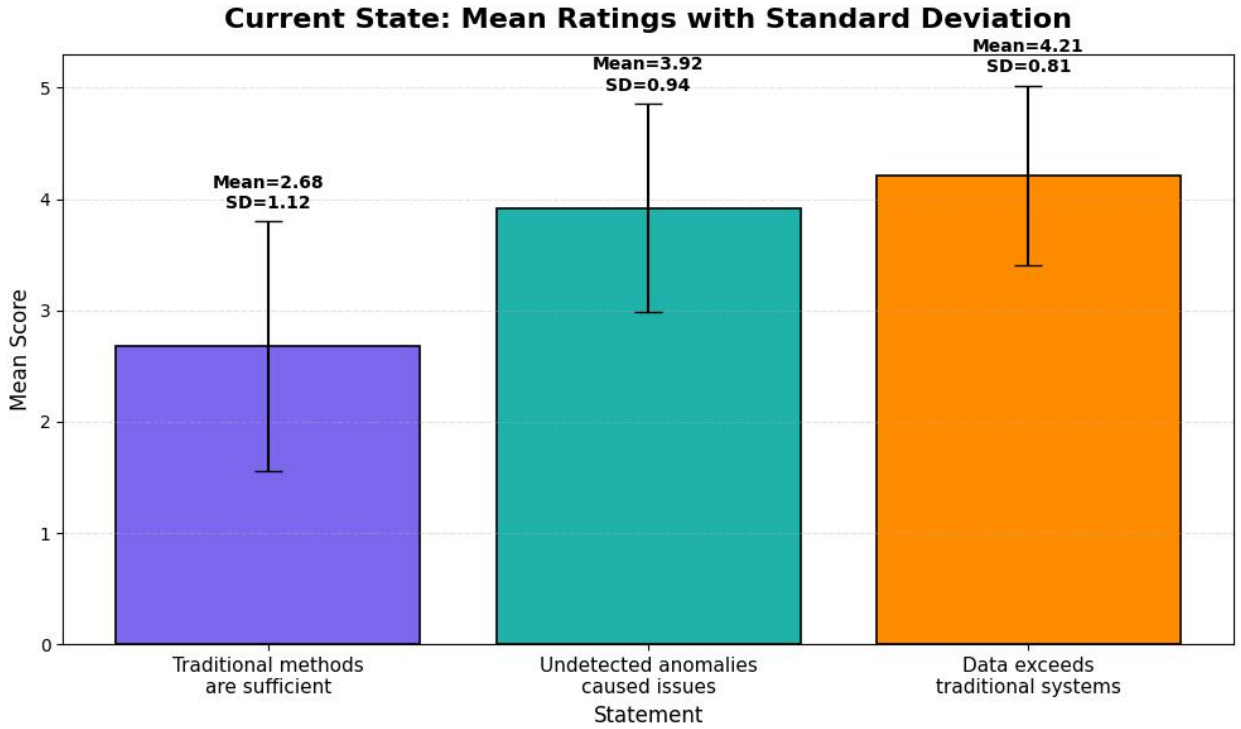

Fig 1: Current State

The findings reveal a definite trend away on the trust in conventional ways. The mean of the statement is low (2.68) and the variability (SD = 1.12) is highest which suggests that the participants do not agree in general with some deviation on opinions. Conversely, respondents confirm that there are problems with undetected anomalies (Mean = 3.92, SD = 0.94), which points to the fact that the current strategies are not always effective in detecting problems.

The greatest agreement is found in the case of Data exceeds traditional systems (Mean = 4.21, SD = 0.81) as a wide majority of the participants agree that current data volumes and complexity exceed the ability of traditional solutions.

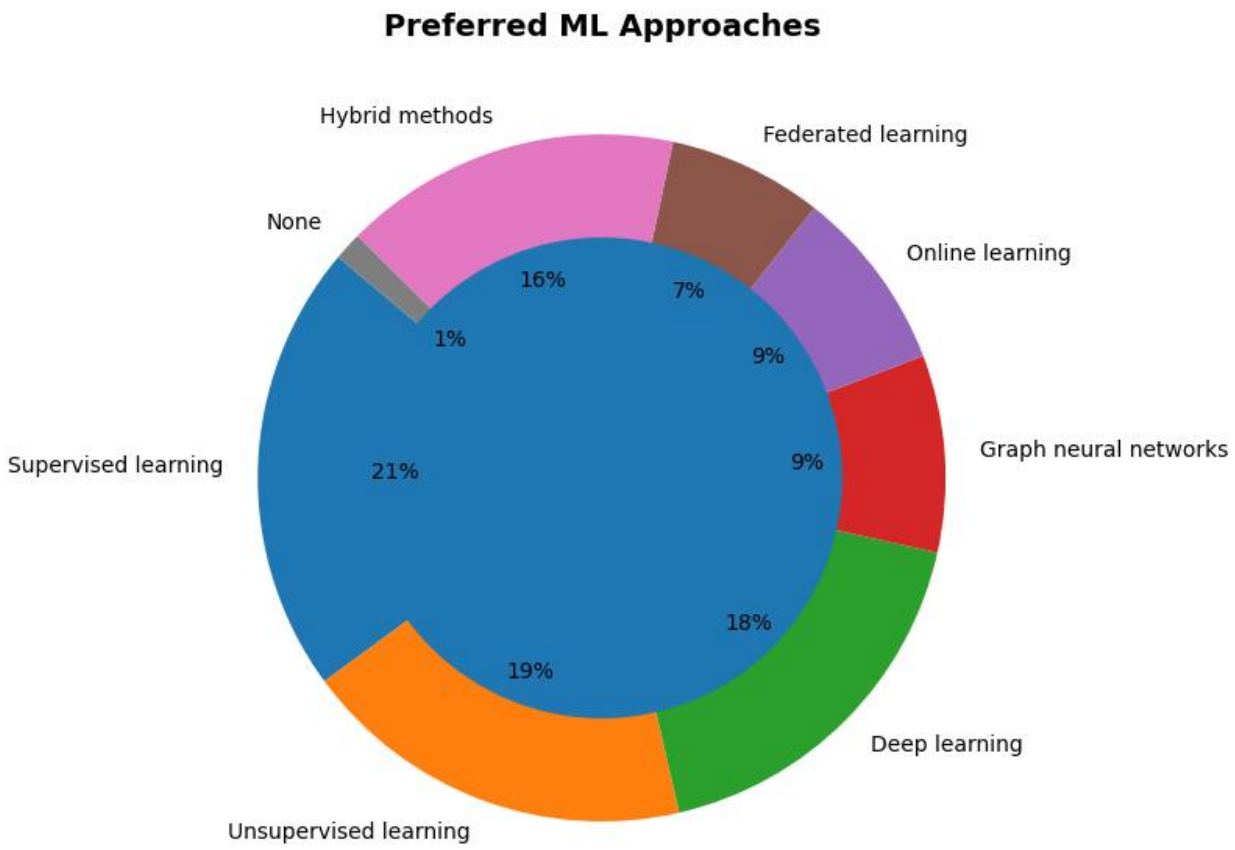

Fig 2: Preferred ML Approaches

Results indicate that there is a high preference towards existing and generalized methods of machine learning. The most preferred method is supervised learning (64%), which implies using labeled data and demonstrated predictive accuracy. Unsupervised learning (56%) and deep learning (54%), also have a strong support, as there is a need to operate with complex patterns and unlabeled data.

Hybrid techniques (48%) are rather high, which could indicate an increasing trend of using several techniques to enhance robustness and performance. Conversely, more advanced or recent methods like graph neural networks (28%), online learning (26%), and federated learning (22%) are moderately used, probably because of greater complexity of implementation or less generalizability.

The minimal selection of "None" (4%) confirms broad acceptance of ML-based solutions overall.

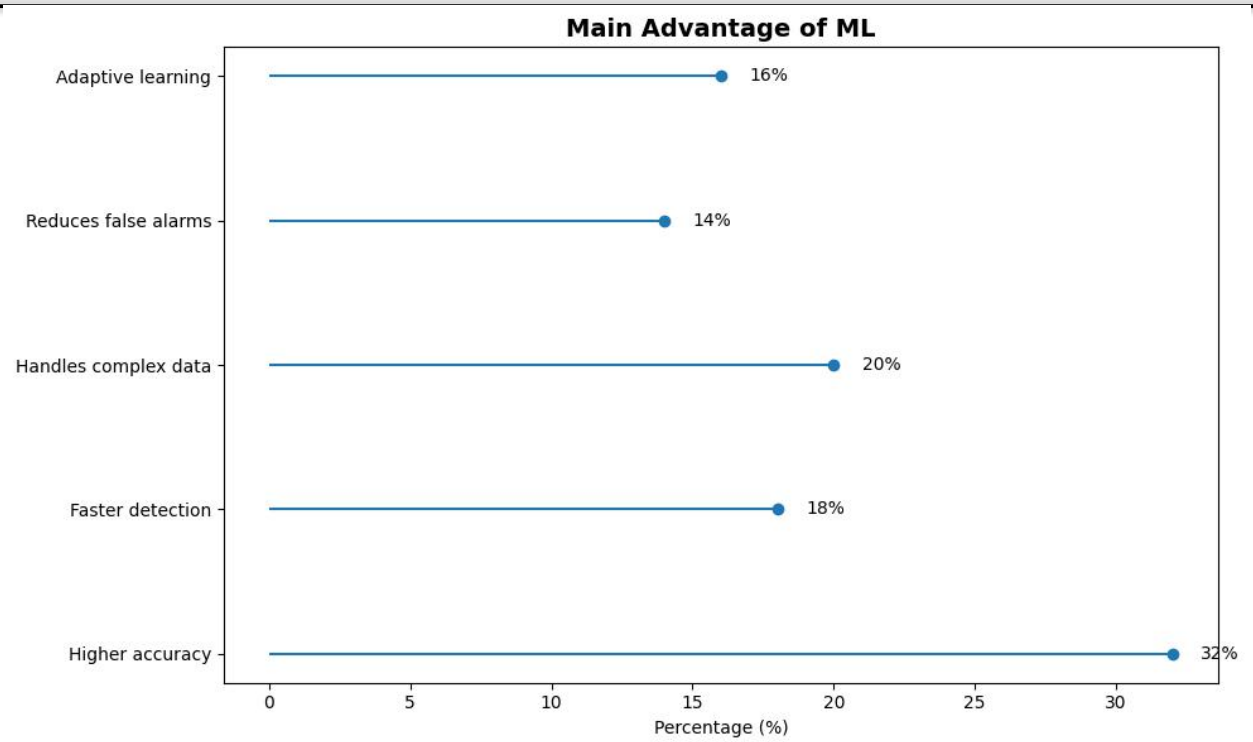

Fig 3: Main Advantage of ML

The findings reveal that the main benefit of machine learning is considered to be higher accuracy (32%), which is its key strength to enhance the reliability of the decision. This is then followed by working with complex data (20%), which supports the use of ML in the processing of massive and high-dimensional data that conventional data processing techniques cannot deal with. Moderate attention is given to faster detection (18%) and adaptive learning (16), with the real-time performance and the capability to learn with developing data being valued but not a priority. The lowest perceived benefits are reduction in false alarms (14%), which denotes that it is viewed as an advantage but not the strongest contributor.

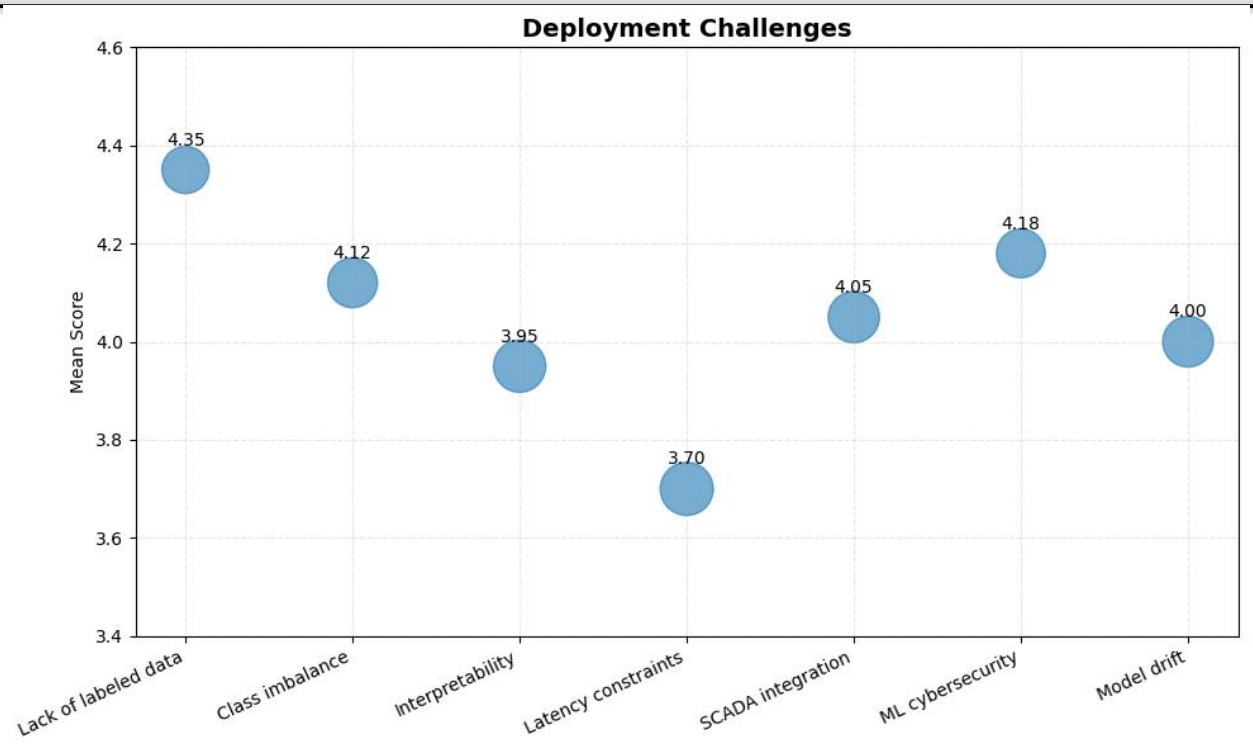

Fig 4: Deployment Challenges

The findings emphasize data-related limitations as the most significant obstacles to deploying ML. The mean of lack of labeled data is the greatest (4.35, SD = 0.72), which means that there is a high level of agreement around the main challenge which is a lack of data. Likewise, the imbalance by classes (Mean = 4.12, SD = 0.80) adds to the issues of data quality and distribution.

The security and system integration become also significant concerns. ML cybersecurity (Mean = 4.18, SD = 0.77) displays great concern about vulnerability whereas SCADA integration (Mean = 4.05, SD = 0.85) indicates a pragmatic challenge in integrating ML into the current operational infrastructures.

The challenges of the model are of medium, yet significant concern. Model drift (Mean = 4.00, SD = 0.83) and interpretability (Mean = 3.95, SD = 0.88) implies the need of maintainable and explainable systems. Latency constraints (Mean = 3.70, SD = 0.91) is the lowest ranked which is still above neutral meaning that it is an issue but not the most important.

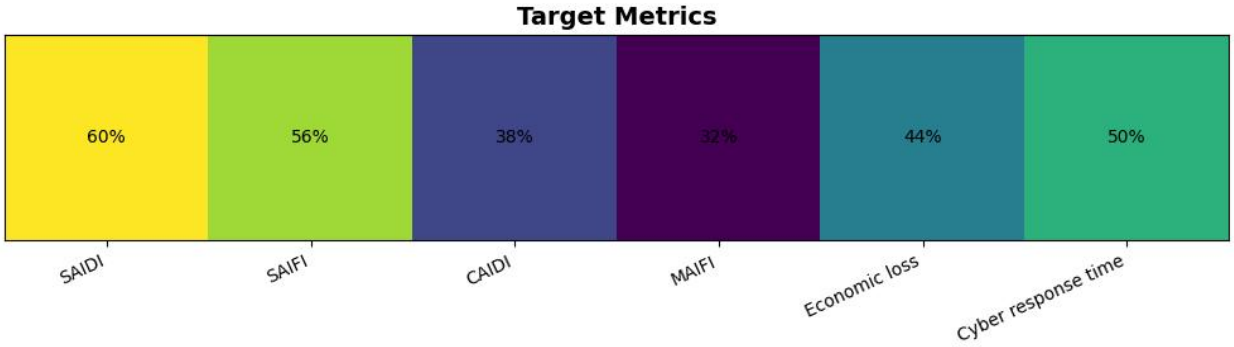

Fig 5: Target Metrices

The findings indicate that attention was paid to performance metrics of reliability and outage to a large degree. The highest priority is given to SAIDI (60%) and SAIFI (56%) meaning that the shortest time and the decreased frequency of outage is the main goal of respondents.

Cyber response time (50%) is also on the top list, which indicates the growing interest in quick detection and mitigation of cyber threats and the connection of cybersecurity with operational performance outcomes. Economic loss (44%) is next, indicating that the financial consideration is also significant and secondary.

The importance of CAIDI (38%), MAIFI (32%) receives less focus, which means that although these aspects are also relevant, they are not the priority.

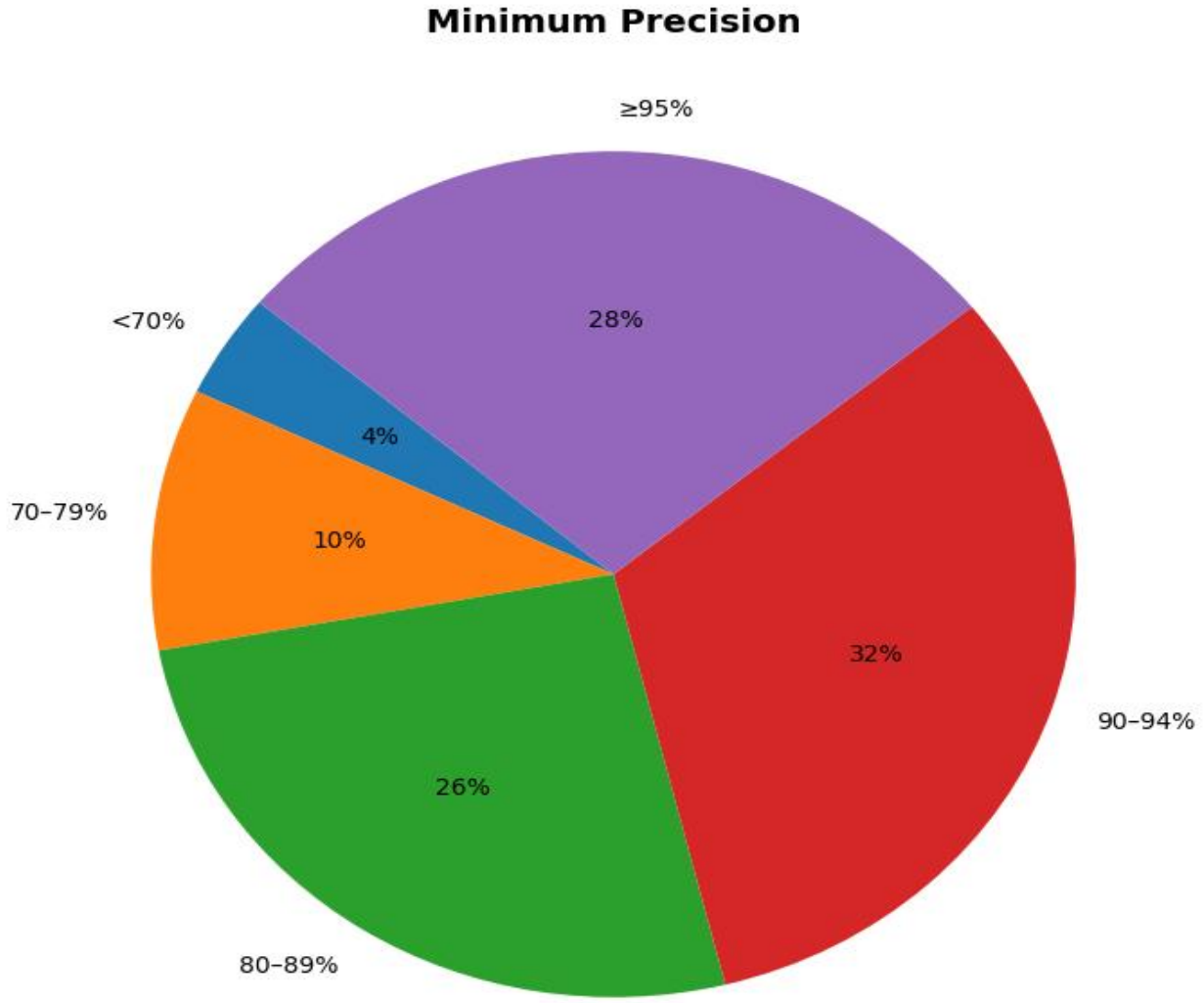

Fig 6: Minimum Precision

The distribution shows that high-precision ML models are highly favored. Most respondents anticipate a precision level of 90% or above with 90 to 94% (32%) and ≥95% (28%) being used together by 60% of respondents. This indicates a great need in very efficient model performance.

The moderate percentage (26%) regards 80-89% as acceptable and implies that there is a tolerance of slightly less precision in less critical uses.

Nonetheless, few respondents are willing to take lower thresholds with only 10% of them opting to take 70-79% and only 4% indicating <70%.

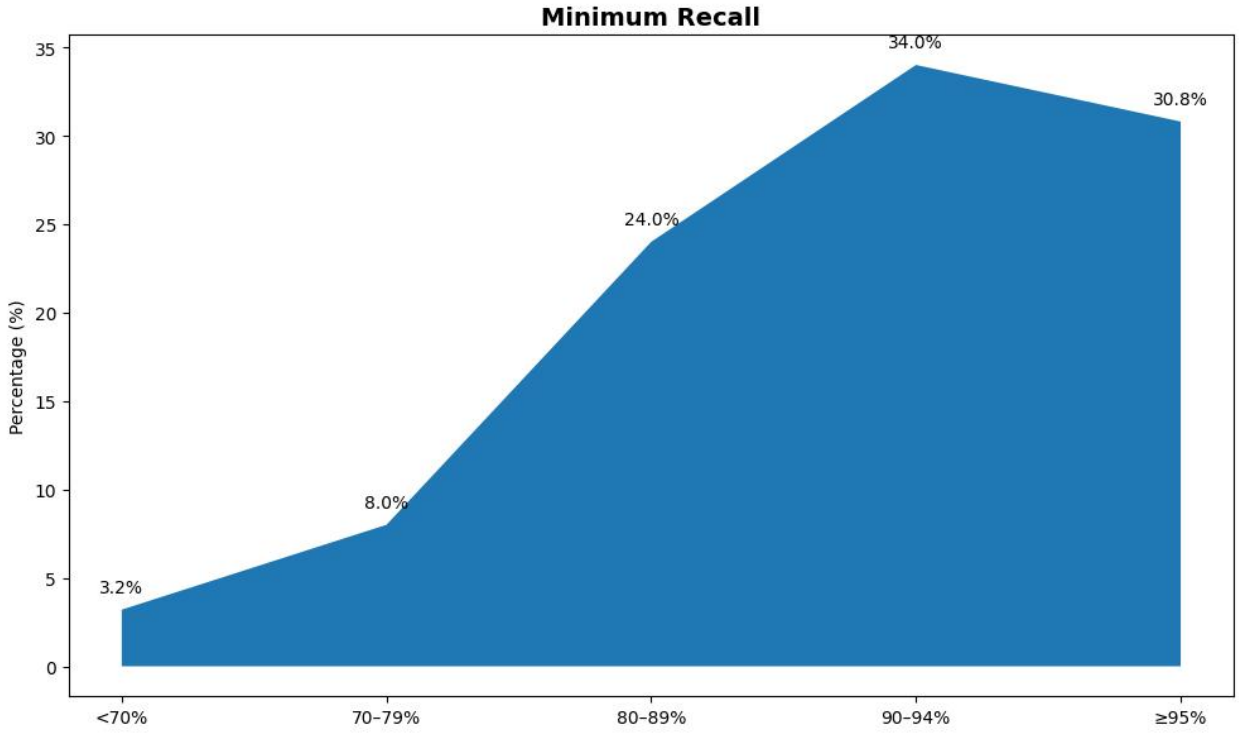

Fig 7: Minimum Recall

The findings show that there is a high expectation of high recall which means that minimizing missed detections are very necessary. A clear majority of respondents (64.8%) require recall levels of 90% or higher, with 90–94% (34%) being the most selected range, followed by ≥95% (30.8%). Another significant segment (24%) tolerates 80 - 89%, indicating that moderately high recall is not extensively flexible enough in low-stakes situations. However, tolerance for lower recall is minimal, with only 8% selecting 70–79% and 3.2% indicating <70%.

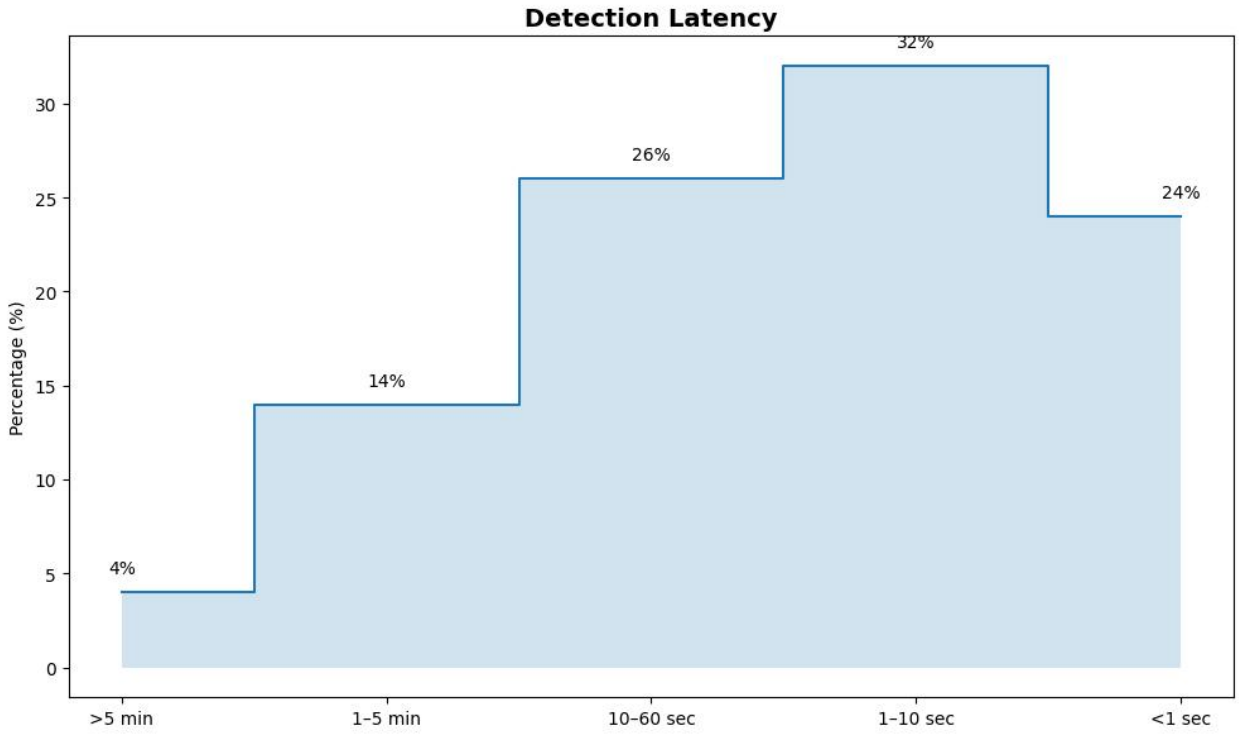

Fig 8: Detection Latency

The findings show that there is a high demand towards near real-time detection features. The majority of respondents (56%) expect latency within 10 seconds, with 1–10 sec (32%) being the most preferred range, followed by <1 sec (24%). This underscores the importance of the urgent reaction within working settings.

The remaining 26% tolerate 10-60 seconds which implies that there is certain leniency to short delays during which an immediate reaction is not as important. However, longer latencies are largely unacceptable, with only 14% selecting 1–5 minutes and just 4% indicating >5 minutes.

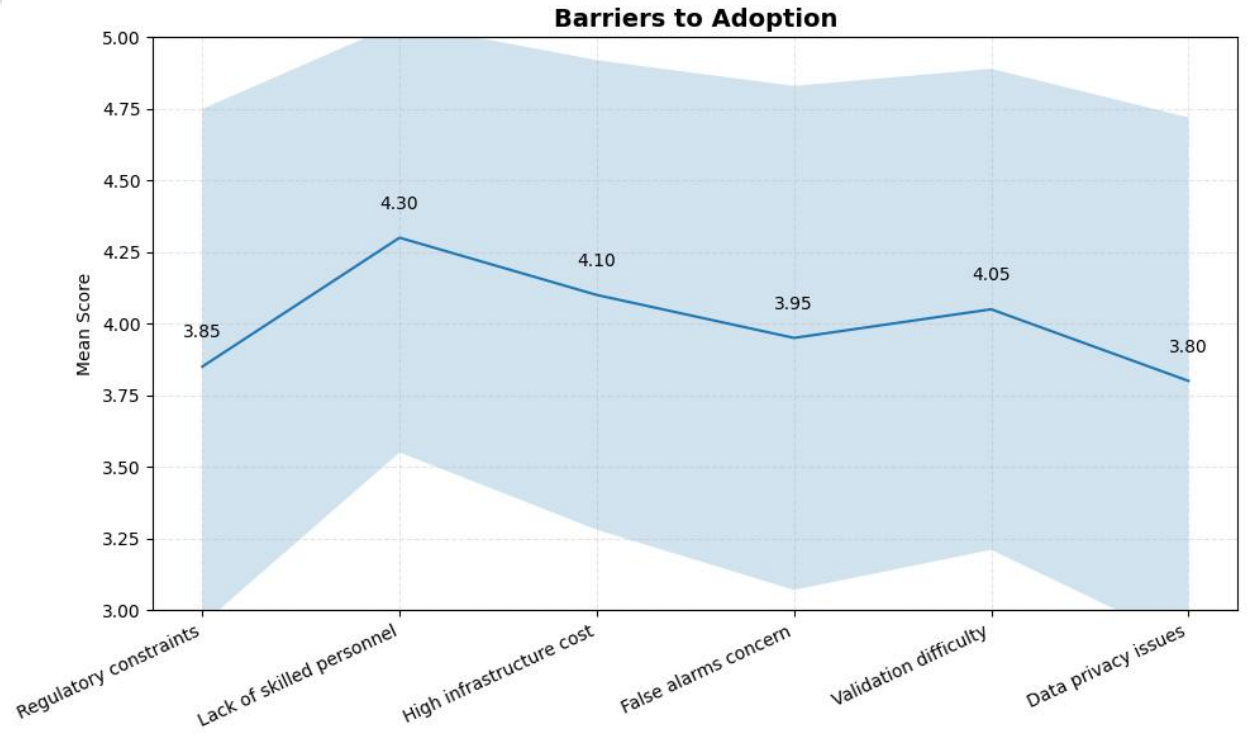

Fig 9: Barriers to Adoption

The results show that the most notable of these are human and resource constraints. The mean of lack of skilled personnel is the highest (4.30, SD = 0.75) with a high and consistent level of agreement that expertise gaps are a barrier to adoption. The cost of infrastructure is also noted to be a barrier to the implementation and testing of the ML systems (High infrastructure cost Mean = 4.10, SD = 0.82) and the validation is also a barrier (High infrastructure cost Mean = 4.05, SD = 0.84). There are also operational issues. False alarms (Mean = 3.95, SD = 0.88) indicate fears of reliability of the system and trust. In the meantime, regulatory constraints (Mean = 3.85, SD = 0.90) and data privacy issues (Mean = 3.80, SD = 0.92) are moderately concerning but slightly more variable, as it demonstrates differences in the perceptions of respondents.

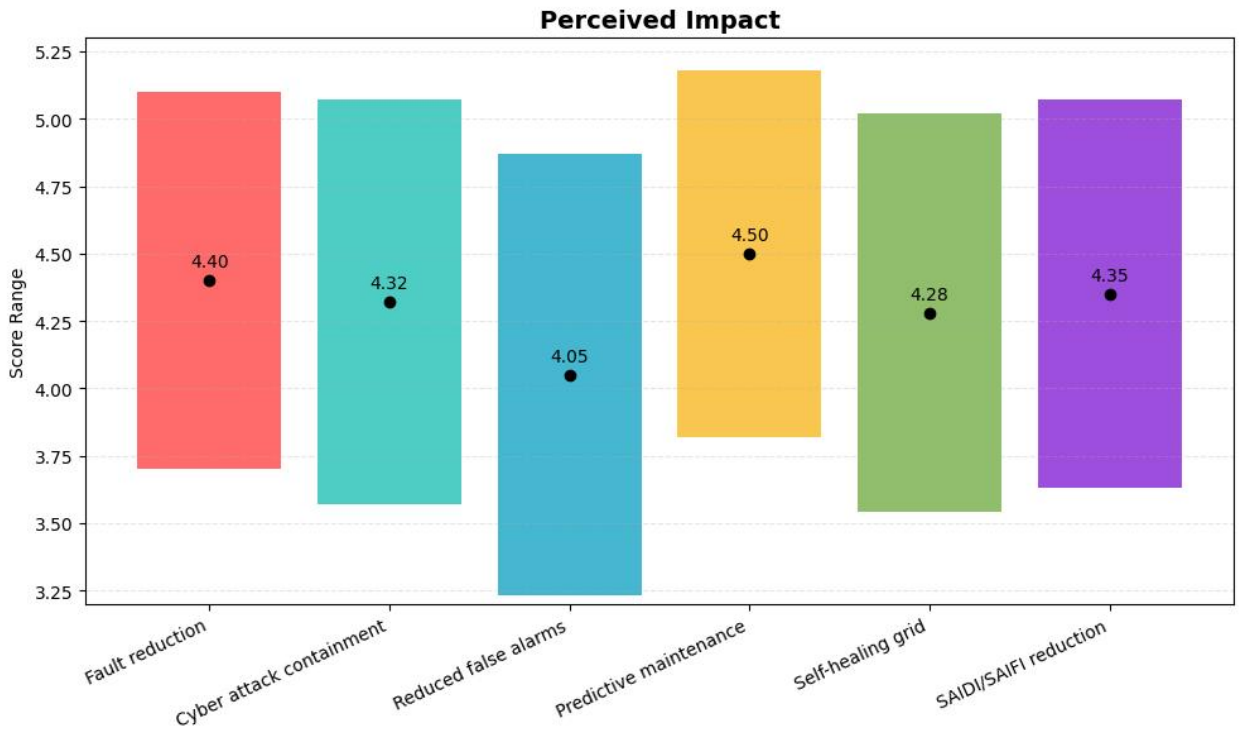

Fig 10: Perceived Impact

The findings indicate that there is a strong agreement as the perceived value of ML is high on all of the impact areas with the mean score exceeding 4.0.

Predictive maintenance is the most influential (Mean = 4.50, SD = 0.68) indicating that it is perceived as a highly influential application, especially in regard to failure prevention and optimization of the assets. Fault reduction (4.40) and SAIDI/SAIFI reduction (4.35) come right behind, further supporting the idea of enhancing system reliability and reducing outages.

Containment of cyber attacks (4.32) and self-healing grids (4.28) are also rated high, which demonstrates trust in ML to raise resilience and automation levels in systems. False alarms reduced (4.05) are a bit lower but also positive. The standard deviations are relatively low in all aspects indicating high agreement amongst the respondents.

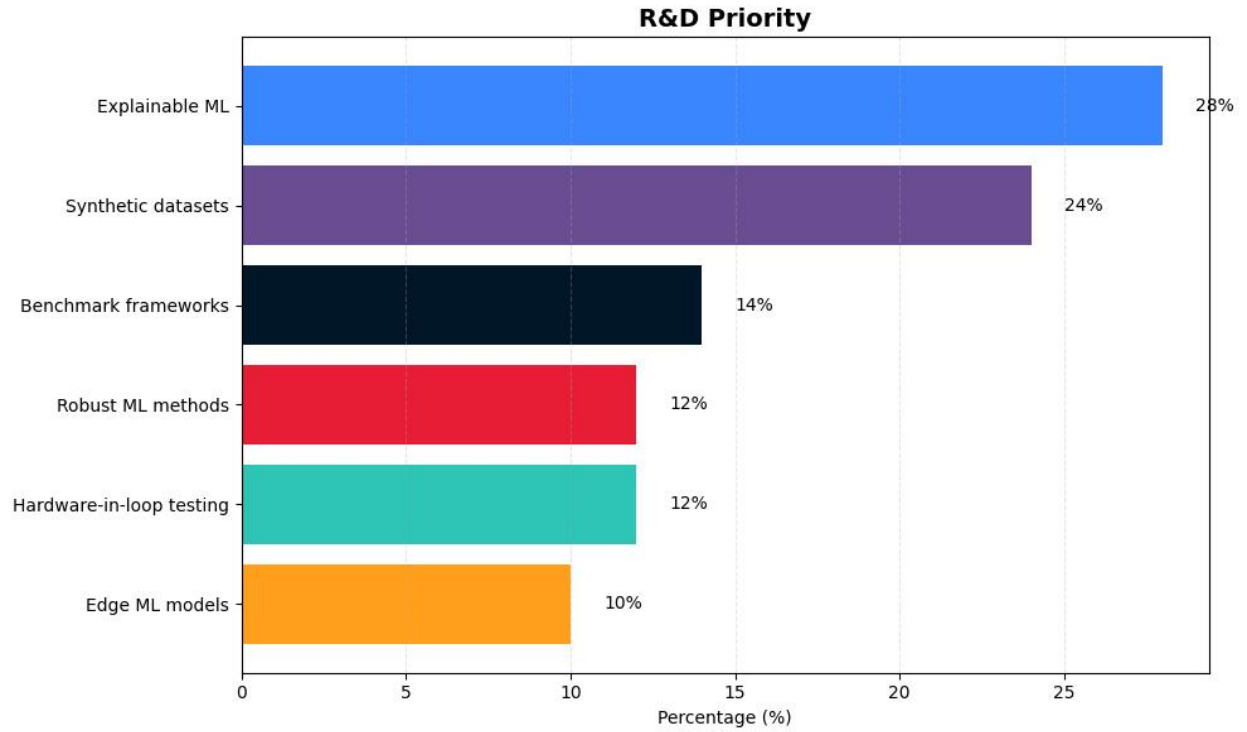

Fig 11: R&D Priority

Our insights reveal that the explainability of the model and the data are clearly at the forefront of research needs. The top one is explainable ML (28%) which shows the need for understanding and trust in ML systems, particularly in critical uses of ML systems. Followed by synthetic data (24%), where the need to address the absence of data and privacy concerns are evident.

The benchmark frameworks (14%), the need for standardized evaluation, and two hardware-in-loop testing (12%) and robust ML methods (12%), real-world validation and stability issues are all mid-level priorities.

Compared to more foundational issues like explainability and data, edge ML models (10) are given a relatively lower priority, suggesting that edge deployment is significant yet not as just as pressing as more basic assets.

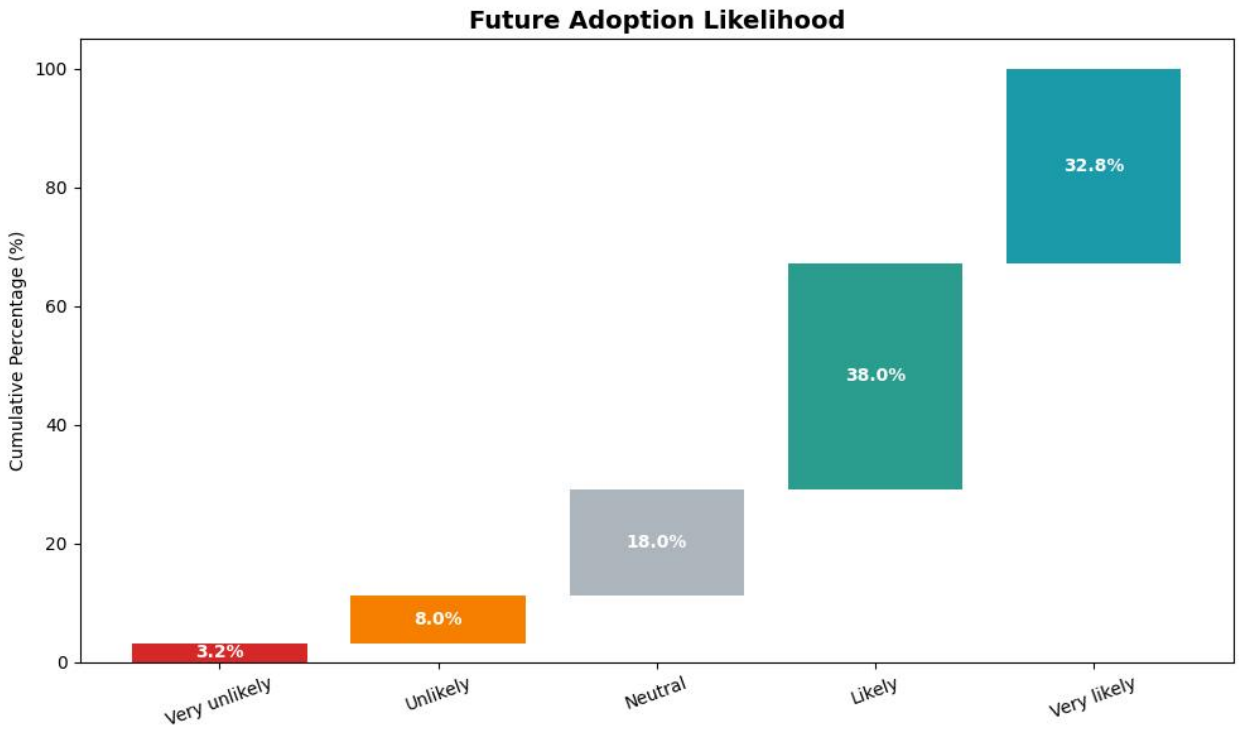

Fig 12: Future Adoption Likelihood

The outcomes of the survey show that there is a bright future for the use of ML. The majority (70.8%) of the respondents are positive with the most common responses to Likely (38%) and very likely (32.8%). Hence, optimism towards the use of ML.

A moderate proportion (18%) is Neutral which can be interpreted as uncertain or wait and see until proven more mature or not or until price comes down. There is a low optimist (8% Unlikely and 3.2% Very unlikely), so there are not much negative views.

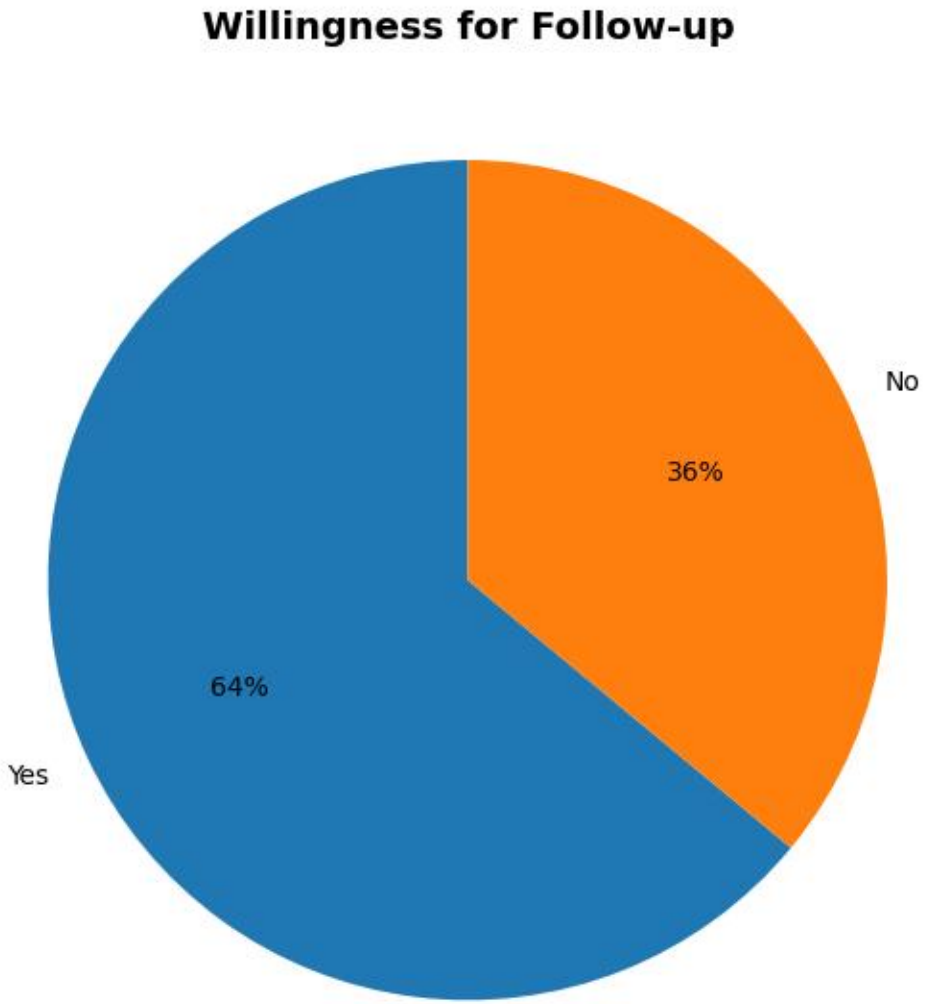

Fig 13: Willingness for Follow-up

There is considerable interaction and willingness to participate. A large proportion (64%) answered Yes, which suggests the willingness to engage further, provide feedback or co-operate.

However, a substantial number (36%) responded No, suggesting that while general participation is desirable, many of these participants can only participate to a limited extent because of constraints such as time, lack of incentives or less interest.

Discussion

The results of this research support the recent acknowledgement that traditional anomaly detection systems are not appropriate for smart grids. The dissatisfaction with traditional approaches is supported by literature reporting that traditional methods are inefficient in dealing with large amounts of high-speed data for anomaly detection [4], The high level of agreement that the data is too complex for traditional systems also points to the need for newer data-driven methods as observed in recent literature [3], [5].

The high preference for supervised, unsupervised and deep learning methods indicates a balance between predictive accuracy and adaptability. This accords with the research that shows that supervised learning methods have high accuracy when data is labelled, while unsupervised and deep learning methods capture sparse and dynamic anomalies [21], [23]. The strong interest in hybrid approaches also suggests a shift towards more robust and generalised approaches, as is acknowledged in research that shows how ensemble approaches increase detection rates for smart grids [26].

While the benefits are recognised, the findings highlight the challenges posed in deploying ML, including lack of labeled data and class imbalances. These are known to affect training and performance [27]. Other themes including cybersecurity and SCADA integration address issues related to ML deployment into a smart grid environment in line with previous research on cyber-physical systems integration and cybersecurity threats [12], [28].

The performance targets outlined in our research - high accuracy and recall (90%) and low detection time (10 seconds) - underline the importance of smart grid functions. This performance level matches the requirements for real-time anomaly detection in other research where the ability to detect problems quickly and with high accuracy is critical for safety and reliability [29]. The emphasis on service reliability indices (SAIDI and SAIFI) also suggests the aim is to enhance reliability through rapid diagnosis and isolation of faults.

The barriers to adopting (such as lack of staffing and equipment costs) indicate that non-technical (human and economic) factors still play a major role. This finding is in line with previous studies that technical readiness should be complemented by human resources and resources [11]. In addition, the investment in research and development on explainable ML and synthetic data generation suggests a need for confidence and data, which are both critical for uptake.

Finally, the high impact and positive expected adoption suggest there are high expectations for ML to improve smart grid reliability. The potential use cases such as predictive maintenance and cyber mitigation control are highly promising, and are in line with other research that has demonstrated the use of ML for active and resilient grid operation [30], [31]. In conclusion, despite the challenges, it appears that the outlook for uptake and integration of machine learning in smart grids is promising.

Conclusion and Recommendations

This research shows that, as grids have become smart, they have become more complex, and that traditional anomaly detection techniques will not keep them reliable. The findings demonstrate that experts agree that machine learning (ML) can offer an improvement, particularly for large and high-dimensional data and dynamic and complex anomalies. The strong preference (supervised, unsupervised and hybrid) for ML methods highlights the need for both accuracy and adaptability in real-world applications. Further, the emphasis on high precision, recall and low latency highlights the importance of anomaly detection in smart grids.

Despite acknowledging the advantages, the research shows that ML is still in its infancy, with adoption at the pilot and partial level across most organisations. Some of the major obstacles to deployment, including the absence of labeled data, class imbalance, integration with existing systems, and cybersecurity issues, remain. Moreover, internal factors such as lack of skilled resources and investment in equipment and infrastructure pose additional limitations. However, the significant perceived benefits of ML for predictive maintenance, fault mitigation and cyber-attack prevention, and the very positive outlook for future adoption, suggest a positive future for integration.

Based on these findings, several recommendations are proposed. First, data strategies such as generating synthetic data and better labeling of data should be focused on to address data scarcity and class imbalance. Second, focus should be given to explainable AI (XAI) to increase transparency and trust among grid operators and other stakeholders. Third, human capacity building is needed to build capacity for development and implementation.

In terms of the technical, consideration should be given to incorporating ML algorithms into SCADA and grid management systems for interoperability and scalability. Also, real-time testing and low-latency system designs should be explored. Improving the security of ML-based systems should also be considered.

Finally, governments and industry leaders should promote interdisciplinary research and standards development, including benchmarking and hardware-in-the-loop testing, to promote innovation and adoption. Addressing these technical, operational and regulatory challenges will allow ML anomaly detection to transform smart grid reliability and resilience.

References

[1] J. Singh, O. A. Shah, and S. Arora, "Smart grid cybersecurity: Anomaly detection in solar power systems using deep learning," *Energy Storage and Saving*, 2025.

[2] N. Sahani, R. Zhu, J. H. Cho, and C. C. Liu, "Machine learning-based intrusion detection for smart grid computing: A survey," *ACM Trans. Cyber-Phys. Syst.*, vol. 7, no. 2, pp. 1–31, 2023.

[3] C. Xu, P. Zhang, N. Luo, F. Zheng, and W. He, "Integrating machine learning for anomaly detection and pattern recognition in smart grid power data," *Distrib. Gener. Altern. Energy J.*, pp. 595–614, 2025.

[4] M. O. Rahaman and M. R. Mahin, "Machine learning techniques for anomaly detection in smart grids," *J. Humanit. Soc. Sci. Stud.*, vol. 6, no. 12, pp. 159–165, 2024.

[5] M. S. Abdalzaher, M. F. Shaaban, and R. Aburukba, "Leveraging optimized machine learning for anomaly detection and quality of service enhancement of PMUs observation in smart grids," *IEEE Access*, vol. 13, pp. 196959–196972, 2025.

[6] E. Omol, L. Mburu, and D. Onyango, "Anomaly detection in IoT sensor data using machine learning techniques for predictive maintenance in smart grids," *Int. J. Sci. Technol. Manag.*, vol. 5, no. 1, pp. 201–210, 2024.

[7] X. J. Li, M. Ma, and Y. Sun, "An adaptive deep learning neural network model to enhance machine-learning-based classifiers for intrusion detection in smart grids," *Algorithms*, vol. 16, no. 6, p. 288, 2023.

[8] A. Shaheen, "Cybersecurity in the modern era: An overview of recent trends," *J. Eng. Comput. Intell. Rev.*, vol. 1, no. 1, pp. 39–50, 2023.

[9] M. T. Abbas and F. Hanif, "Design and control of an autonomous drone navigation system using embedded AI," *J. Eng. Comput. Intell. Rev.*, vol. 3, no. 2, pp. 68–80, 2025.

[10] U. Imtiaz *et al.*, "An integrated machine learning framework for structural health monitoring of bridges: A case study on Soan Bridge," *Asian Bull. Big Data Manag.*, vol. 5, no. 2, pp. 194–207, 2025.

[11] F. Amin, N. Daudpota, and A. Khan, "A complete penetration testing framework: Simulating attacks and evaluating post-exploitation techniques with Kali Linux and Metasploit," *Spectrum Eng. Sci.*, pp. 386–407, 2025.

[12] U. Imtiaz, S. Malik, and A. Khan, "Blockchain-driven cybersecurity framework for smart homes: Integrating IoT and machine learning for secure automation," *Asian Bull. Big Data Manag.*, vol. 4, no. 4, pp. 570–583, 2024.

[13] F. Amin, "Binary flaw detection: A security analysis paper," in *Proc. Int. Conf. Adv. Mach. Intell. Cybersecurity Technol. (AMICT)*, Kota Kinabalu, Malaysia, 2025, pp. 325–330, doi: 10.1109/AMICT65811.2025.11402666.

[14] T. A. Shiva, N. Ireen, and M. S. Islam, "Optimizing early intervention strategies for neurodiverse children (ASD): Reducing long-term public healthcare costs through parent-mediated training," *Apex J. Soc. Sci.*, vol. 3, no. 1, pp. 30–52, 2024.

[15] M. S. Islam and T. A. Shiva, "Virtual cognitive behavioural therapy in rural US communities: Effectiveness and reach," *J. Bus. Insight Innov.*, vol. 3, no. 2, pp. 60–76, 2024.

[16] U. Twaha, A. Mosaddeque, and M. Rowshon, "Accounting implications of using AI to enhance incentives for wireless energy transmission in smart cities," *Int. J. Manag. Res. Glob. Educ.*, vol. 6, no. 2, pp. 1208–1218, 2025.

[17] M. K. Khan and A. Ullah, "Implication of IoT and its impact on library services: An overview," *Inverge J. Soc. Sci.*, vol. 3, no. 2, pp. 63–72, 2024.

[18] M. M. S. Nevisi *et al.*, "An evolutionary deep reinforcement learning-based framework for efficient anomaly detection in smart power distribution grids," *Energies*, vol. 18, no. 10, p. 2435, 2025.

[19] J. Wang, "Multi-agent system based smart grid anomaly detection using blockchain

machine learning model in mobile edge computing network," *Comput. Electr. Eng.*, vol. 121, p. 109825, 2025.

[20] S. Abbas *et al.*, "Improving smart grids security: An active learning approach for smart grid-based energy theft detection," *IEEE Access*, vol. 12, pp. 1706–1717, 2023.

[21] K. Alam, M. Al Imran, U. Mahmud, and A. Al Fathah, "Cyber attacks detection and mitigation using machine learning in smart grid systems," *J. Sci. Eng. Res.*, p. 12, 2024.

[22] R. Shrestha *et al.*, "Anomaly detection based on LSTM and autoencoders using federated learning in smart electric grid," *J. Parallel Distrib. Comput.*, vol. 193, p. 104951, 2024.

[23] A. Aljohani, M. AlMuhaini, H. V. Poor, and H. M. Binqadhi, "A deep learning-based cyber intrusion detection and mitigation system for smart grids," *IEEE Trans. Artif. Intell.*, vol. 5, no. 8, pp. 3902–3914, 2024.

[24] S. M. Shakil, A. Hossain, and M. M. Rahman, "An empirical evaluation of anomaly detection techniques in smart grid systems using real-time operational data," *Am. J. Adv. Technol. Eng. Sol.*, vol. 1, no. 2, pp. 95–134, 2025.

[25] V. K. Singh and M. Govindarasu, "A cyber-physical anomaly detection for wide-area protection using machine learning," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3514–3526, 2021.

[26] R. Qi, C. Rasband, J. Zheng, and R. Longoria, "Detecting cyber attacks in smart grids using semi-supervised anomaly detection and deep representation learning," *Information*, vol. 12, no. 8, p. 328, 2021.

[27] S. Ness, "Adversarial attack detection in smart grids using deep learning architectures," *IEEE Access*, vol. 13, pp. 16314–16323, 2024.

[28] T. S. Ustun *et al.*, "Machine learning-based intrusion detection for achieving cybersecurity in smart grids using IEC 61850 GOOSE messages," *Symmetry*, vol. 13, no. 5, p. 826, 2021.

[29] L. Mascali *et al.*, "A machine learning-based anomaly detection framework for building electricity consumption data," *Sustain. Energy, Grids Netw.*, vol. 36, p. 101194, 2023.

[30] A. Fathollahi, "Machine learning and artificial intelligence techniques in smart grids stability analysis: A review," *Energies*, vol. 18, no. 13, p. 3431, 2025.

[31] H. Nhung-Nguyen *et al.*, "Machine-learning-based anomaly detection for GOOSE in digital substations," *Energies*, vol. 17, no. 15, p. 3745, 2024.